Automatic human face detection for home surveillance application

Citation for published version (APA):

Document status and date:
Published: 01/01/2002

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Theorem 3. The scheme described in this section is an unconditionally secure proactive secret sharing scheme if the following conditions are satisfied:

1. \( \Gamma = 3 \times \Gamma \),
2. \( \Im(M_{\phi}(G)) = K^d \) if and only if \( G \in \Gamma \). (i.e., \( M_{\phi}(G) \) has a maximal column rank if and only if \( G \in \Gamma \).

The proof follows from Theorem 2 and from the description of the protocol.

References

2. Face detection. When the person gets near enough, his/her face can be detected and extracted. The corresponding facial features are extracted and/or coded at the same time for later processing.

3. Face recognition. Face recognition distinguishes different persons and decides about the person's identification.

Among these tasks, face detection is an indispensable and important step because it usually sets the input for the face recognition modules. The performance of most face recognition algorithms relies on an accurate locating of human face regions. However, face detection has remained to be a difficult problem because human face varies strongly in its appearance. The challenges associated with face detection can be attributed to the following factors: pose, presence or absence of structural components, facial expression, occlusion, image orientation and imaging conditions [3].

A wide variety of methods for face detection have been proposed in the past decade. Extended surveys of these methods can be found in [3][4]. These methods range from simple edge-based algorithms to complex high-level approaches utilizing advanced pattern recognition techniques. Among these approaches, skin color has been proved to be an effective facial feature for locating facial regions. In addition, skin-color-based methods usually achieve much higher speed than other computation-expensive approaches and are especially suitable for real-time applications such as video surveillance. Several studies [1][2][3] have revealed that human skin color forms a highly condensed cluster in certain color spaces, and its distribution can be characterized as a multivariate Gaussian distribution. However, skin color detection methods are prone to detection errors, especially in cases where human faces are immersed in a skin-color-like background. Some recent skin-color based methods use geometric analysis or motion information for face verification. However, these methods generally cannot achieve robustness due to the following reasons:

1. Most verification processes only simply make a true-or-false decision for a candidate skin-color blob, therefore, they are prone to make erroneous decisions when faces and their background are contained in the same skin-color region.
2. In geometry-based verification, a set of rules defining a face are usually formed according to heuristics or anthropometric measurements. Due to the variation of human faces (e.g. race, sex, and age), these rules should be well defined. The definition of a 'perfect' threshold for classification of faces and non-faces has been proved not feasible.

In this paper, we propose a skin-color based face detection which invokes a further extraction process. The extraction process uses confidence values for some facial features. The confidence value of each feature is based on a probabilistic analysis of facial geometry. The proposed approach selects the most probable facial feature combination to form a complete facial region. Our approach has the following advantages over aforementioned skin-color-based methods:

1. It can extract the face region from a large skin-colored area by verifying the existence of facial features such as mouth and eyes.
2. It avoids a strict true-or-false decision by incorporating probabilistic metrics. A set of facial features is considered to have a certain probability to represent a human face. If one set of facial features achieve a sufficiently high confidence value, the corresponding region is regarded as a real face region.

The paper is structured as follows. In section 2, we give a brief introduction of the proposed method and its processing flow. Section 3 illustrates in more detail the key algorithms, i.e. the skin-color segmentation and probability-based facial region evaluation. Section 4 presents the evaluation results of experiments using our proposal. Section 5 presents conclusions.

2 Sequence of Processing Algorithms for Face Detection

In this section, we present an overview of all processing tasks that are performed for our face detection proposal. The purpose of this section is to briefly discuss the individual steps in the processing. In the next section, we will discuss details of the most important steps. Fig. 1 visualizes the sequence of processing steps.
After these pre-processing steps, we use a probability-based face verification algorithm to further extract facial regions from the candidate regions. In this algorithm, possible prominent facial feature regions (mouth and eyes) are first located, using histogram analysis. Then so-called individual local confidence values are assigned to each feature. All possible combinations of facial features are then formed to generate a face candidate set. For each combination, the global relationship between constituting feature regions is examined to yield a global confidence value. The candidate combination with the largest comprehensive confidence value is considered the most-likely facial region. After such a facial region is determined, post-processing eliminates unlikely candidates based on shape analysis (Fig. 1(g)). The rectangular bounding box indicates the face region, while the enclosed upper and lower horizontal lines represent the eye region and mouth region, respectively.

3 Key Algorithms for Face Detection

3.1 Skin Color Segmentation

We apply the research result from [11] because it uses two signal components only which supports real-time applications. We analyzed a sample set of facial pixels taken from people of different race, gender, age and under various lighting conditions. The skin color is mostly concentrated in two normalized signal components \( r = \frac{R}{R + G + B} \) and \( g = \frac{G}{R + G + B} \). The combination \((r, g)\) yields a bivariate Gaussian distribution, which is plotted in Fig. 2.

![Fig. 2. Bivariate Gaussian distribution of skin color showing a narrow concentration](image)

The mean values of \( r \) and \( g \) \((\mu_r = 0.4157, \mu_g = 0.4125, \sigma_r^2 = 154.9733, \sigma_g^2 = 20.1298, \sigma_{rg} = -18.5518)\) are computed. Also the covariance term \( \sigma_{rg} \) and the signal variances \( \sigma_r^2 \) and \( \sigma_g^2 \) are evaluated for a rough segmentation of the original image frame into possible facial areas.

3.2 Probability-based Facial Region Verification

Facial features have been widely used in facial processing literature for classifying human faces. However, facial feature appearance and distribution vary considerably among different people and under different conditions. Some of the facial features cannot be explicitly distinguished from image deteriorations resulting from local poor video quality (e.g. a dark noisy area). Some of the facial features can be easily mixed up like eyebrows and eyes. For this reason, we focus on first selecting the most prominent facial features - eyes and mouth. We assume that these features are always present in every face image. At the same time, the detection of less prominent features, like nostrils and eyebrows, can contribute positively to a confident identification of prominent facial features.

The algorithm starts with a binary image which may contain a human face. First the horizontal histogram is generated from the input image [6], from which second derivatives of local maxima are located as reference lines. These reference lines correspond to some comparatively 'flat' areas having no significant dark regions, i.e. potential cheek lines. The local minima above the reference line and the local minima below the reference line form two candidate sets: the eye region candidate set \( ES_k = \{e_i \mid i = 1, m\} \) and the mouth region candidate set \( MS_k = \{m_j \mid j = 1, n\} \) (see Fig. 3). The algorithm iterates over the whole facial region to get all reference lines and all possible eye and mouth candidates: \( ES = \bigcup ES_k \) and \( MS = \bigcup MS_k \).

![Fig. 3. Face detection and subsequent facial histogram analysis](image)

For each \( e_i \in ES \), we define the probability of being the real eye region as a local confidence function:

\[
LC(e_i, \text{EYE}) = \sum_{a=1}^{m} w_a P_a .
\]

The term \( w_a \) is a weighting factor and \( P_a \) represents the confidence value based on different assessment criteria that are explained below. Two criteria are of significant importance for their contribution to the confidence function.

1. A principal criterion is the matching probability of \( e_i \) fitting to a pre-defined eye region template (see Fig. 3(a)). A series of templates are used in this stage to accommodate for various eye positions. The probability value indicates how well \( e_i \) fits to one of these available templates.
Another important criterion considers the possible existence of less prominent features like nostrils and eyebrows, for increasing the reliability of our decision. The detection of less prominent features in the correct positions relative to the eyes increases the probability that an auxiliary region is a real eye region. In Fig. 3(b), auxiliary regions \( e_i \) and \( e_k \) are introduced for this purpose, where \( e_i \) may contain the eyebrows and \( e_k \) may contain the cheek. The presence of low-intensity pixels in \( e_i \) and high-intensity pixels in \( e_k \) gives \( e_i \) a higher confidence value. This is especially useful when a significant eyebrow area appears in the face.

![Eye templates](image)

(a) Fig. 4. Local confidence evaluation

Similar processing is also applied to evaluate the mouth region. After local confidence evaluation is examined, the individual candidate features of \( ES \) and \( MS \) are extended with their corresponding confidence values, as follows:

\[
ESC = \{(e_i, c_{e_i}) \mid e_i \in ES\}, \quad MSC = \{(m_j, c_{m_j}) \mid m_j \in MS\},
\]

where \( c_{e_i} \) represents the confidence value of \( e_i \) and \( c_{m_j} \) represents the confidence value of \( m_j \).

The examination of only local features and their relations is not sufficient for a final judgment of a facial region. A global confidence evaluation process is applied to assess the overall probability for a face candidate. We define a face candidate as a combination of facial feature regions. For each face candidate \( (e_i, m_j) \in ESC \times MSC \), a global confidence function is evaluated as follows:

\[
GC(e_i, m_j) = \sum_{t=1}^{n} w_t Q_t.
\]

The term \( w_t \) is a weighting factor and \( Q_t \) represents the global confidence value according to global assessment criteria. These criteria are based on geometric relations between different features. For example, one of the criteria is based on the statistical analysis of the ratio between face width at the eye level \( d_y \) and the distance between eye and mouth \( d_m \). The probability density function of this parameter for a large group of people can be characterized as a normal distribution. Therefore, the confidence value \( Q_t \) in our proposal is based on the probability evaluation of the specific ratio \( d_y/d_m \).

The final confidence evaluation function of a face candidate combines equations (1) and (2) into:

\[
Confidence(e_i, m_j) = w_y GC(e_i, m_j) + w_e LC(e_i) + w_m LC(m_j).
\]

The variables \( w_y, w_e, w_m \) are weighting factors. The face candidate with the highest confidence value is selected as the most probable facial region. A post-processing step is then performed to make a final judgment using a shape analysis of the selected region.

4 Experimental Results

To evaluate the efficiency of our proposed technique, we tested our method on several video sequences including both standard testing sequences and self-made sequences. The system achieves an average of 10.7 frames/sec processing speed on a Pentium IV desktop (Intel 1.7G Pentium IV processor, 256M memory, Red Hat Linux 7.2, gcc) with a resolution of 320 pixels by 240 lines, which is sufficient for real-time video surveillance applications.

Example frames from the testing video sequences are shown in Fig. 5. In sequence Salesman (Fig. 5(a)) which contains relatively detailed background, the system identifies the face region correctly in 90% of the frames. We recorded a special experimental indoor sequence (Fig. 5(b)) because it contains skin-coloured background objects such as cabinets, tables and parts of the person's clothes. The system successfully identifies the face region in 95% of the frames. The experiments prove that our approach is quite attractive for real-time video surveillance applications.

However, initially the proposed method was prone to make false detections when dealing with more detailed background objects, and it showed low robustness against poorly-illuminated environments. These problems were solved by incorporating an image correction pre-processing step and a refinement of the previously mentioned confidence evaluation criteria.
5 Conclusions

In this paper, we have proposed a fast human face detection technique for home video surveillance applications. The proposal first uses a skin-color model to roughly segment the possible facial areas. Afterwards, a probability-based evaluation process is applied to further extract facial features from candidate regions. In this step, prominent feature regions such as eyes and mouth are extracted based on confidence value evaluation. The experimental results show that a 90\% detection accuracy is achieved. As there is no training procedure involved in the processing, the proposed technique is suitable for real-time video processing.

Our proposal can be further refined by incorporating more advanced normalization and criterion selection. Motion information can also be included to improve the overall detection accuracy. The robustness of the system with respect to poor illumination and rotation of the faces still need further study.

References


Multi-party Server-Aided Key Distribution Protocols Based on Symmetric Techniques

Malek Rechlaghem*
Université Libre de Bruxelles
Dept. of Computer Science, Bd. Du Triomphe, C.P. 212.
1050 Brussels, Belgium
mrechlag@ulb.ac.be

Abstract. Server-aided multi-party key distribution (MSKD) protocols are protocols in which a trusted Third Party generates and broadcasts securely a session key to protocol participants. We investigate MSKD protocols through their desired properties and propose new protocols offering additional properties. Our protocols are contributory in the sense that every protocol participant equally contributes to the key generation process and guarantees its freshness. Our protocols do not require the use of conventional encryption/decryption, instead, we use other security mechanisms such as Secure Keyed Hash Functions. Our protocols allow pre-computation on TTP side which is an important property allowing to reduce the TTP computational complexity during protocol runs. We define backward secrecy as being a weak variant of perfect forward secrecy and show that our protocol provide backward secrecy.

I. Introduction

Key establishment protocols can be roughly classified in two categories: Key distribution protocols which rely on a trusted entity to generate and transfer a session key to protocol participants. On the other hand, in Key agreement protocols each participant is involved in the construction of the key.

In this paper, we are interested in multi-party key distribution protocols based on symmetric techniques. Server-less multi-party key distribution protocols [5] are not easily extendible to multi-party settings since every user needs to share his secret key with all other users so in this paper we confine our work on Multi-party Server-aided Key Distribution (MSKD) protocols. In this paper, we considered relevant features and properties of MSKD protocols and in doing so, we discovered that existing protocols suffer from the lack several of these properties which led us to define new useful properties.

The remainder of this paper is organized as follows. After presenting some necessary definitions in section II, we discuss the motivations and goals of our work in section III. After describing few useful notations in section IV, we then detail our protocols in section V.

* This paper was finalized when the author was working for Belgacom, Bd du Roi Albert II 27, 1030 Brussels.